TRINITY COLLEGE

Year 12 Mathematics Specialist Units 3, 4
Test 6 2020

Calculator Assumed
Statistical Inference

STUDENT'S NAME Solutions

DATE: Thursday 10 September **TIME:** 50 minutes **MARKS:** 45

INSTRUCTIONS:

Standard Items: Pens, pencils, drawing templates, eraser

Special Items: Three calculators, notes on one side of a single A4 page (these notes to be handed in with this assessment)

Questions or parts of questions worth more than 2 marks require working to be shown to receive full marks.

---

1. (3 marks)

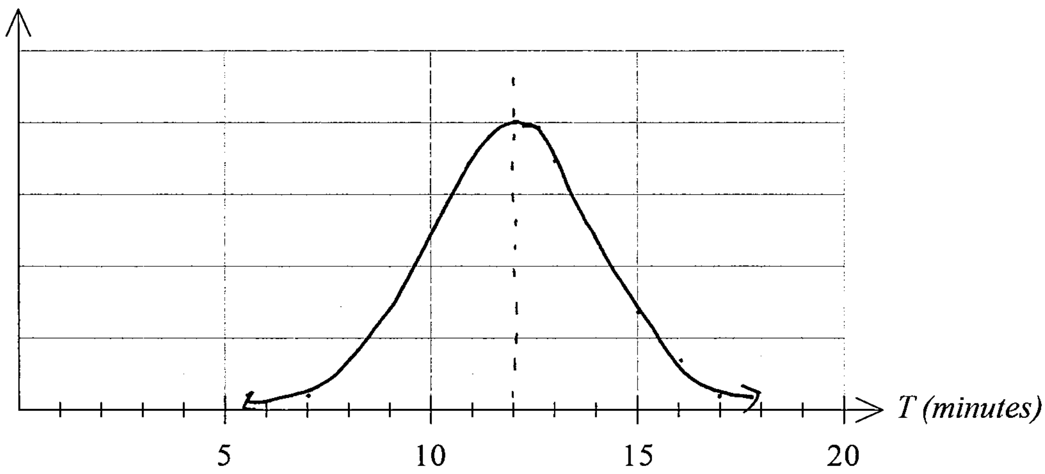
The time T in minutes that it takes a taxi to arrive is uniformly distributed with a population mean of $\mu(T) = 12$ and a population variance of $\sigma^2(T) = 81$.

If a large number of samples, each with a sample size of 16 times is taken, sketch the likely distribution of the sample mean $\bar{T}$ below.

$\sigma = 9$ $\quad \sigma_{\bar{x}} = \frac{9}{\sqrt{16}} = \frac{9}{4}$ ✓ $\quad \mu_{\bar{x}} = 12$

✓ mean

✓ ~95% in $7 \le \mu \le 17$



2. (6 marks)

Consider the following statement:

The 95% confidence interval for the mean ranges from 0.1 to 0.4.

Determine, with reasons, if the following statements are TRUE or FALSE.

(a) *The probability that the true mean is greater than 0 is at least 95%.* [2]

False.

We cannot infer any probability from a confidence interval

(b) *There is a 95% probability that the true mean lies between 0.1 and 0.4.* [2]

False.

We cannot infer probability from a confidence interval. The confidence interval either contains the parameter or not – it does not contain the parameter with any degree of probability.

(c) *If we were to repeat the experiment over and over, then 95% of the time the true mean falls between 0.1 and 0.4.* [2]

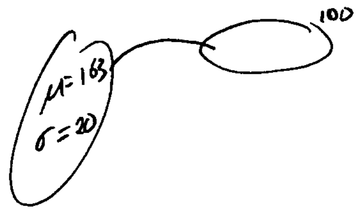
False.

Every time a random sample is taken, the measurements are different, so the boundaries of the confidence interval change.

3. (10 marks)

100 cricket balls are weighed. It is known that the population mean $\mu = 163\ g$ and the population standard deviation $\sigma = 20\ g$

(a) State the (approximate) distribution of the sample mean weight per cricket ball. [3]

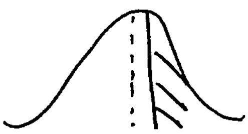


$n \geq 30$, so approx normal ✓

$\bar{X} \sim N\left(163, \left(\frac{20}{\sqrt{100}}\right)^2\right)$

✓ ✓

(b) Determine the probability that the sample mean weight of a cricket ball will be more than $163.5g$. [2]



$P(\bar{X} \geq 163.5) = 0.4013$

Suppose that more than 100 cricket balls are weighed.

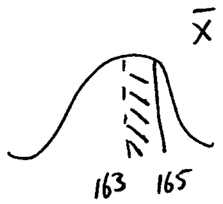
(c) How would this affect your answer to part (b)? Explain without recalculation. [2]

As you increase the sample size, the st dev of $\bar{x}$ would decrease.

$\therefore$ $P(\bar{X} \geq 163.5)$ would decrease.

The sports club desires that the probability that the sample mean will be between $163g$ and $165g$ is greater than 45%.

(d) Determine the minimum number of cricket balls that will need to be weighed. [3]



$P(163 \leq \bar{X} \leq 165) \geq 0.45$

$\Rightarrow \sigma_{\bar{X}} = 1.2159$

$\therefore 1.2159 = \frac{20}{\sqrt{n}}$

$n = 270.55$

$\hat{=} 271$

OR $P(0 \leq z \leq k) \geq 0.45$

$k > 1.6449$

Now $\frac{x - \mu}{\sigma/\sqrt{n}} = z$

$\frac{165 - 163}{20/\sqrt{n}} = 1.6449$

$n \approx 271$

4. (9 marks)

Matthew and Angus want to check the claim that a single Kitkat weights 17 g. They buy two packets that contain 18 Kitkats each, and weigh them both for a combined weight of 620 g. Matthew calculates the sample standard deviation to be $s = 1.2\ g$ and assumes that the weights of the Kitkats is uniformly distributed.

(a) Based on Matthew's data, obtain a 95% confidence interval for μ, the population mean weight of a Kitkat. [4]

$\bar{x}=17.2$, $s=1.2$

$$\bar{x} = \frac{620}{36} = 17.\dot{2}$$

$n \geq 30$, approx normal.

95% C.I. $\quad z = 1.96$

$\therefore$ 95% C.I. $\quad 16.83 \leq \mu \leq 17.61$

Angus points out that the population distribution is not uniformly distributed but is normally distributed.

(b) How does this affect the confidence interval calculated in part (a)? [2]

No change. Distribution of sample means will be normal due to C.L.T.

A different sample of 36 Kitkats is taken and it is found that the sample standard deviation is $s = 0.9\ g$. A confidence interval for the population mean weight is determined to be $16.92 \leq \mu \leq 17.48$.

(c) Determine the level of confidence, to the nearest 0.01%, used to calculate this interval. [3]

$$\bar{x} = \frac{17.48 + 16.92}{2} = 17.2$$

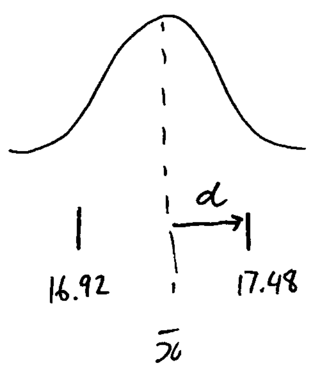
$$d = 17.48 - 17.2 = 0.28$$

Now $d = z\dfrac{s}{\sqrt{n}}$

$$0.28 = z \times \frac{0.9}{\sqrt{36}}$$

$$z = 1.8667$$

Now $P(-1.8667 \leq z \leq 1.8667) = 0.9381$



$\therefore$ 93.81% C.I.

5. (9 marks)

The cable for the new zip line on the Matagarup bridge is required to support a weight of 12 kN. A first random sample of n cables found the mean cable strength to be 12.5 kN. Repeated sampling of the mean indicated that the standard deviation of the sample means was 300 N.

(a) Determine a 90% confidence interval for the cable strength mean, correct to the nearest N. [3]

90% C.I. $\Rightarrow$ $z = 1.6449$

$12.5 - 1.65 \times 0.3 \le \mu \le 12.5 + 1.65 \times 0.3$

$12.007 \le \mu \le 12.993$

A second random sample of $4n$ cables found that the average cable strength was 12.3 kN. Assume that both samples were drawn from the same population.

(b) What is the standard deviation of the sample mean for the second sample, current to the nearest N [2]

$\frac{\sigma_1}{\sqrt{n}} = 300$ $\qquad$ $\sigma_2 = \frac{\sigma}{\sqrt{4n}} = \frac{1}{2}\frac{\sigma}{\sqrt{n}}$ $\qquad$ $\therefore \sigma_2 = \frac{1}{2} \times 300 = 150$ N

or 0.15 kN

Suppose that the first random sample and the second random sample are combined to produce a third random sample of $5n$ cables. Consider the 90% confidence intervals as:

Random Sample	***Size***	***90% Confidence interval***
First	n	A
Second	$4n$	B
Third	$5n$	C

(c) Which of the intervals, A, B or C, will provide the greatest precision in determining the population mean μ? Justify your answer. [2]

C, it has the largest sample size, so therefore the smallest standard error.

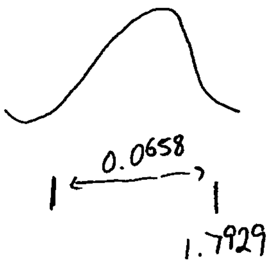
(d) Which of the intervals, A, B or C, contains the true value of the population mean μ? Justify your answer. [2]

We don't know. Random sampling is random. We may or may not have the population mean in each interval.

6. (8 marks)

Declan wants to estimate the population mean height in metres, of the current year 12 cohort. He takes a random sample of 50 students and determines a 99% confidence interval for μ. The upper limit of this interval is 1.7929 m and the width of this interval is 0.0658 m.

(a) Determine the sample mean for this sample of 50 students. [2]



$$\bar{x} = 1.7929 - \frac{0.0658}{2}$$

$$= 1.76 \text{ m}$$

(b) Calculate, correct to 0.01 metres, the sample standard deviation for the sample of 50 students. [3]

99% C.I. $\Rightarrow$ $z = 2.5758$

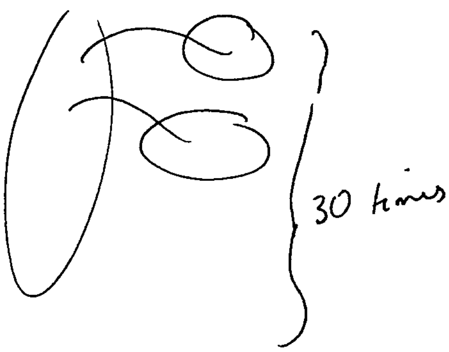
Now $d = \frac{zs}{\sqrt{n}}$

$$\frac{0.0658}{2} = \frac{2.5758\, s}{\sqrt{50}} \quad \Rightarrow \quad s = 0.0903$$

$$\doteq 0.09 \text{ m}$$

Declan knows that taller people have a greater chance to play basketball for the College. He knows from previous experience that in a group of 50 students, there is a 6.3% chance that this group has at least one student who plays basketball for the College with a standard deviation of 1.72. Declan decides to investigate this further and randomly samples groups of 50 students a total of 30 times.

(c) State the approximate distribution of the mean number of samples of 50 students who have at least one student that plays basketball for the College. [3]



$n = 30$, approx normal

$$\therefore \bar{Y} \sim N\left(0.063, \left(\frac{1.72}{\sqrt{30}}\right)^2\right)$$

Y: group with at least 1 basketball player